

## Telephone Traffic Prediction Based on Modified Forecasting Model

<sup>1</sup>Jiangbao Li, <sup>1</sup>Zhenhong Jia, <sup>1</sup>Xizhong Qin, <sup>2</sup>Lei Sheng and <sup>2</sup>Li Chen

<sup>1</sup>School of Information Science and Engineering, Xinjiang University, Urumqi 830046, China

<sup>2</sup>Subsidiary Company of China Mobile in Xinjiang, Urumqi 830063, China

---

**Abstract:** This study presents a busy telephone traffic prediction model that combines wavelet transformation and least squares support vector machine. Firstly, decompose preprocessed telephone traffic data with Mallat algorithm and get low frequency component and high frequency component. Secondly, reconfigure each component and use LS\_SVM model to predict each reconfigure one. Then the traffic can be achieved. The results of experiments have testified higher prediction accuracy and stability of this combined traffic prediction model.

**Keywords:** Busy telephone traffic prediction, least squares support vector machines, wavelet transformation

---

### INTRODUCTION

With the continuously increasing of mobile users, telephone traffic is also increased. And in holidays, mobile communication network are faced with the great influence of heavy telephone traffic. In order to prevent network congestion and improve the utilization ratio of network resources, accurate prediction of the future telephone traffic has an important significance. Traditional linear prediction model of time series include Poisson distribution model, Markov model, AR model and ARMA model (Jie *et al.*, 2007), Poisson model cannot describe the characteristics of the contemporary telephone traffic. Markov model has high complexity of the calculation. The algorithm of AR model and ARMA model is simple, but the prediction error is high. ARIMA (Wang and Shen, 2010), FARIMA (Xiao-Tian and Jing-Xian, 2011) prediction model can better capture long and short data-related characteristics, but for non-linear and non-stationary traffic, they do not forecast accurately. Neural network model (Jun-Song, 2009; Xu-Qi and Zong-Tao, 2011) reflect the inherent law of data through the training, but the neural network is based on empirical risk minimization principle and it is easy to lead to less learning or more learning phenomenon in the training process, so the prediction is not accurate. The method of Support Vector Machine (SVM) is based on structural risk minimization (Cristianini and Shawe-Taylor, 2004; Hai-Shen and Xiao-Ling, 2006; Hua and Hong, 2011; Rui and Zhen-Hong, 2011; Xiao-Dong *et al.*, 2005; Xiao-Qiang and Xian-Min, 2011; Zhi-Hui *et al.*, 2007), it can better solve the data of the small sample, nonlinear, high dimension and other characteristics. The shortcomings of the traditional

SVM is on the issue of quadratic programming, namely, kernel function iteration error accumulation will lead to inaccurate results. And the method use equality constraints of least squares support vector machine (LS\_SVM) to replace the inequality constraints of SVM to solve the quadratic programming problem which is transformed into the problem of linear equations, so as to simplify the model structure and it can improve the training speed. So the LS\_SVM model has been the most widely used. The wavelet transform has strong Multi-scale analysis capability; it can remove the correlation of traffic data. The literature (Jun *et al.*, 2008; Hua-Li and Yuan, 2011; Xiao-Tian and Jing-Xian, 2011) use the combined forecasting model with wavelet, the combination model can achieve good prediction. Single prediction model can't solve various characteristics of telephone traffic, so this study is combined with wavelet prediction model.

For the non-stationary self-similarity and multi-scale characteristic of telephone traffic, this study puts forward a busy telephone traffic prediction model that combines wavelet transformation and least squares support vector machine. Firstly, use the 3-layer wavelet of Mallat algorithm decomposition on the busy traffic data to get the low-frequency component and high-frequency components. Secondly, reconfigure each component and use LS\_SVM model to predict single reconstructed components. Finally the prediction results are superimposed. There must be some errors after LS\_SVM prediction model. These errors in the low frequency part and high frequency part of the prediction may be positive or negative. In the last step for telephone traffic synthesis, it can make positive and negative error offset each other, so the improvement of the wavelet transform can achieve better prediction

results. The simulation experiment shows that this model has higher prediction accuracy and stability.

**TELEPHONE TRAFFIC PREDICTION  
ALGORITHM BASED ON IMPROVED  
WAVELET TRANSFORM AND  
LS-SVM MODEL**

**Algorithm steps:**

- Preprocess the original telephone traffic data, then get  $x'(n)$
- Decompose  $x'(n)$  with three layers wavelet of Mallat algorithm and get low frequency component ( $a_3$ ) and high frequency components ( $d_1, d_2, d_3$ )
- Use Mallat algorithm to reconfigure low frequency component ( $a_3$ ) and high frequency components ( $d_1, d_2, d_3$ ), get  $A_3$  and  $D_1, D_2, D_3$
- Use LS\_SVM model to predict  $A_3, D_1, D_2, D_3$ , get  $A_3', D_1', D_2', D_3'$
- Get the final forecast results:  $x'(n)' = A_3' + D_1' + D_2' + D_3'$

**Data pretreatment:** In mobile communication, there are some changes about the telephone traffic every day and the number of telephone traffic is not the same at different times. Especially in holidays, the busy telephone traffic (peak value) has the largest effect on mobile communication network. Sample telephone traffic one time every hour per day and collect 24 data. Then compare the 24 data and treat the maximum as busy telephone traffic data. Collection of 113 days telephone traffic can get  $113 \times 24$  data and 113 busy data. Decompose 113 busy data and reconfigure each component, we can get  $A_3$  and  $D_1, D_2, D_3$  four components. In consideration of the busy VLR users, VLR boot users, message and system connection rate impacts on traffic, treat the four factors as impact factor of the telephone traffic, then use the four factors and single reconstruction data to compose four new matrixes  $113 \times 5$ . Conduct the front of 112 data as the training sample data to predict the final busy telephone traffic data. If the collection of traffic data is  $x^t(n)$  ( $n \in [1, 113]$  the days,  $t : t \in [0, 23]$  the sampling time every day) and assuming the largest traffic is  $x^{\max}(n)$  per day, then use the busy data of  $X^{\max}(1), X^{\max}(2), \dots, X^{\max}(112)$  to predict the 113<sup>th</sup> day of busy telephone traffic data.

**Wavelet decomposition and reconstruction algorithm:** In 1987, Mallat and some others put forward Mallat algorithm for rapid decomposition and reconstruction. Literature (Hua-Li and Yuan, 2011; Xiao-Tian and Jing-Xian, 2011) use the combination forecast model based on the Mallat algorithm of modified wavelet transform and obtain a better prediction result.

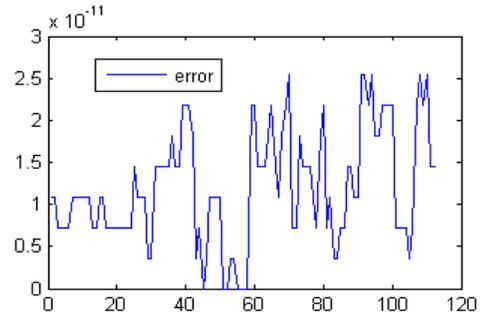


Fig. 1: Data error

The decomposition and reconstruction of Mallat algorithm formula:

$$\begin{cases} a_{j+1} = Ha_j \\ d_{j+1} = Ga_j \end{cases} \quad (1)$$

$$a_{j-1} = a_j \overline{H} + d_j \overline{G}, j = 0, 1, 2, 3, 4 \dots \quad (2)$$

The H is a low-pass filter; G is a high-pass filter. The original signal is decomposed into low frequency part and high frequency part by Mallat algorithm. The low frequency component reflects the outline characteristics and changing tendency of the busy telephone traffic data and high frequency component reflects the impact of dynamic factors such as random disturbance. Low frequency part can be further decomposed, so we can obtain the new low frequency component and high frequency component. Single reconstruction is not to reconstruct the low frequency part and high frequency part at the same time, but separately reconstructed. That's to say when a certain part is reconstructed; the other part should be set to zero.

The difficulty of wavelet decomposition and reconstruction lies in the selection of wavelet basis. This study use different wavelet basis to superpose the reconstruction of the low frequency and high frequency data, then compare with the original data and find that when using biro 1.3 as wavelet basis, the error is  $10^{-11}$  (Fig. 1). While using wavelet basis such as dbN and sym, the error is  $10^{-8}$ . So this study chooses biro 1.3 as wavelet basis.

**The model of LS\_SVM:** The main idea of nonlinear regression of Least Squares Support Vector Machine (LSSVM) is that the input data is mapped to high dimensional feature space H through the nonlinear mapping  $\psi(\cdot)$ , so low dimensional nonlinear regression problem is transformed into linear regression problems in high dimensional feature space. The principle is that we assume that there are 'm' samples data, such as  $(x_1, y_1), (x_2, y_2) \dots (x_m, y_m)$  and soon.  $x_i \in \mathbb{R}^n$  is

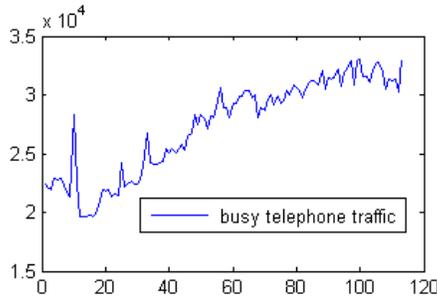


Fig. 2: 113 busy telephone traffic

sample input and  $y_i \in \mathbb{R}$  is the sample output. Construct the optimal regression function:

$$f(x) = w^T \cdot \varphi(x) + b \quad (3)$$

The  $W^T \in \mathbb{H}$ ,  $b \in \mathbb{R}$ ,  $b$  is bias,  $\psi(x)$ : an  $\mathbb{R}^n$ - $\mathbb{H}$  nuclear space mapping function.

According to the principle of solving objective and the structural risk minimization, use Lagrange multiplier method to solve optimization problem. Optimization problem of Lagrange function for least squares support vector machine is:

$$L(w, b, \xi, \alpha) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \xi_i^2 - \sum_{i=1}^m \alpha_i (w \cdot \varphi(x_i) + b + \xi_i - y_i) \quad (4)$$

$C$  = Regularization parameters

$\xi_i$  = The error function

$\alpha_i$  = Lagrange multipliers

$i = 1, 2, 3, \dots, M$

According to the KKT optimization conditions, get the partial derivative from the type (4) and make the result 0. Use Radial Basis kernel Function (RBF) to solve the problem of high dimension calculation. The optimization problem can be further transformed into the solution of linear equations:

$$\begin{bmatrix} 0 & e^T \\ e & \Omega + \frac{1}{2} C * I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix} \quad (5)$$

where,

$e = [1, 1, \dots, 1]$

$\alpha = [\alpha_1, \alpha_2, \dots, \alpha_m]$

$Y = [y_1, y_2, \dots, y_m]$

$\Omega_{ij} = k(x_i, x_j)$

The final nonlinear regression model is:

$$f(x) = \sum_{i=0}^m \alpha_i K(x, x_i) + b \quad (6)$$

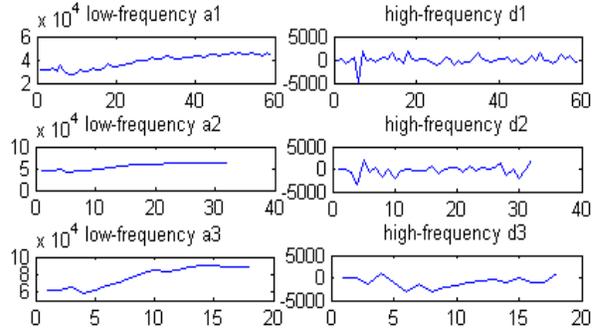


Fig. 3: The decomposition of the busy telephone traffic

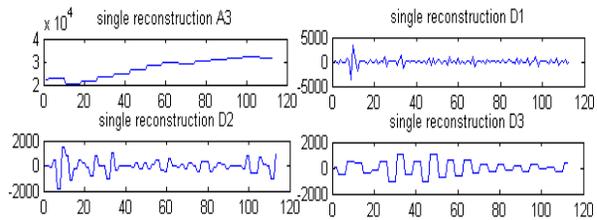


Fig. 4: The single reconstruction of the low frequency component and high frequency component

### THE SIMULATION ANALYSIS

**The data source:** The experimental data is provided by subsidiary company of China mobile in Xinjiang. It Collected the traffic data since January 13<sup>th</sup>, 2012 to May 4<sup>th</sup>, 2012 from Aksu Prefecture of Xinjiang and extracted 113 busy telephone traffic data (Fig. 2). Generally speaking, the more data is selected, the more correctly the learning and training results reflect the relationship between the input and output and the higher precision of prediction is. However, in practical applications, it is impossible to increase the sample data without restrictions, in which case it should try to select representative sample. Considering the correlation between four factors and telephone traffic, the author made some of data pretreatment in this study to guarantee that it would improve the efficiency of the algorithm without losing the original data characteristics. During the May Day, the traffic is larger and on May 4<sup>th</sup> the traffic is largest, this experiment is mainly to predict the busy traffic of May 4<sup>th</sup> day.

**Wavelet decomposition and single reconstruction:** Decompose 113 busy traffic data with Mallat algorithm and get low frequency component and high frequency component (Fig. 3). Then reconfigure each component (Fig. 4). From the results of the decomposition and reconstruction, we can know that the low frequency components conform to the changing tendency of traffic and the high frequency component can reflect the impact of factors such as random disturbance. With the

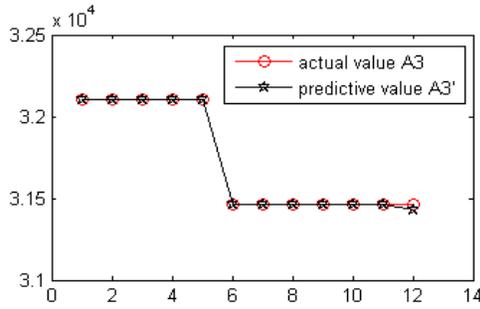


Fig. 5: Prediction of the low-frequency  $A_3$

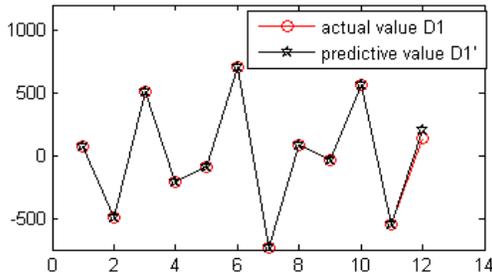


Fig. 6: Prediction of the high-frequency  $D_1$

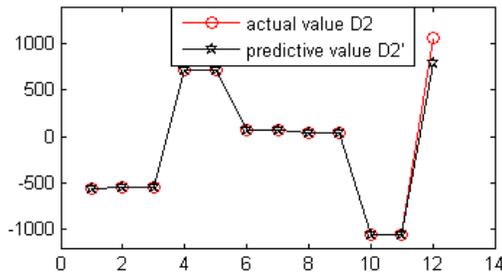


Fig. 7: Prediction of the high-frequency  $D_2$

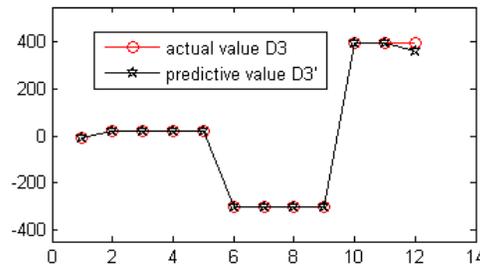


Fig. 8: Prediction of the high-frequency  $D_3$

increasing of the number of decomposition layers, low frequency component and high frequency component become more and more smooth.

Set the busy VLR users, VLR boot users message and system connection rate respectively as  $\eta_1(n)$ ,  $\eta_2(n)$ ,  $\eta_3(n)$  and  $\eta_4(n)$ . When predicting the low frequency information, namely put the  $[\eta_1(n), \eta_2(n), \eta_3(n), \eta_4(n), A_3(n)]$  as input variables ( $n \in [1, 12]$ ), we finally obtain the last day (May 4) of the component

Table 1: The error and prediction efficiency of the prediction model

Times	WT_LSSVM		LS_SVM		Elman	
	Error	Time/s	Error	Time/s	Error	Time/s
1	-0.0103	4.87	-0.0536	1.17	-0.0627	8.07
2	-0.0102	4.75	-0.0536	1.14	-0.0617	8.15
3	-0.0102	4.98	-0.0536	1.18	-0.0642	8.09
4	-0.0103	5.02	-0.0535	1.14	-0.0637	8.13
5	-0.0102	5.02	-0.0534	1.15	-0.0603	8.09
6	-0.0101	4.98	-0.0535	1.15	-0.0614	8.18
7	-0.0102	4.85	-0.0534	1.19	-0.0628	8.20
8	-0.0103	4.97	-0.0535	1.17	-0.0643	8.17
9	-0.0102	4.83	-0.0535	1.14	-0.0631	8.14
10	-0.0101	4.83	-0.0535	1.14	-0.0629	8.16
Mean	-0.0102	4.91	-0.0535	1.16	-0.0627	8.14

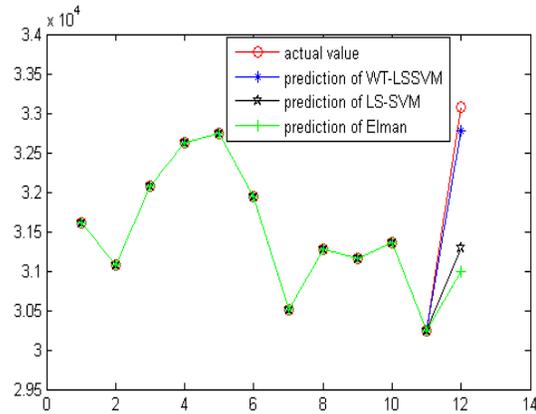


Fig. 9: The comparison of the prediction results

predictive value. The method to predict high frequency components is similar to the low frequency. Selecting different canonical and kernel function parameters for four components prediction can improve prediction accuracy. Low frequency component ( $A_3$ ) and the high frequency component ( $D_1$ ,  $D_2$  and  $D_3$ ) are predicted respectively as shown in Fig. 5 to 8. And then the predicted value of the respective components will be superimposed to obtain the prediction busy telephone traffic of the last day.

**Results analysis:** The method of analyzing relative error is used in this study. The error formula is  $\xi = (x_i' - x_i) / x_i$  ( $x_i'$  is prediction and  $x_i$  is practical value). If the busy telephone traffic on May 4 is 'S' and low frequency component ( $A_3$ ) and the high frequency component ( $D_1$ ,  $D_2$ ,  $D_3$ ) are predicted respectively as  $A_3'$ ,  $D_1'$ ,  $D_2'$  and  $D_3'$ . Then the relative error of WT\_LSSVM prediction model is  $\xi = [(A_3' + D_1' + D_2' + D_3') - S] / S$ . This study chooses LS\_SVM and Elman neural network prediction model as contrast model. The results of prediction shown as Fig. 9. Experiments were conducted 10 times in this study. The error and prediction efficiency of the prediction model for the statistics are shown as in Table 1.

As can be seen, when considering the four impact factors, the prediction of combination forecasting method of WT\_LSSVM is closer to the actual value. This is mainly due to the error of the prediction result

of the respective components which is positive or negative. After superposing each component, the prediction error can be canceled from each other. So the prediction error can be reduced to -0.0102 and it takes 4.91s. The efficiency of LS\_SVM prediction model which only takes 1.16s is higher than the prediction of WT\_LSSVM, but the accuracy of LS\_SVM prediction model is lower, because the relative error is -0.0535. The prediction error of Elman is -0.0627 and it takes 8.14s. Both the precision and efficiency is lower than WT\_LSSVM prediction method. The stability of Elman model is worse than the compared model.

### CONCLUSION

This study put forward a kind of improved telephone traffic measurement algorithm based on wavelet transform and LSSVM model. At first, decompose preprocessed telephone traffic data with Mallat algorithm and get low frequency component and high frequency component, then reconfigure each component and use LS\_SVM model to predict single reconstructed components. In the last step for telephone traffic synthesis, it can make positive and negative error offset each other, so the improvement of the wavelet transform model can achieve better prediction results and more stable. It will start to research prediction model which combines wavelet transform with optimized LSSVM in next step, improving the prediction accuracy further.

### ACKNOWLEDGMENT

First of all, I'd like to express my thankfulness to the judges for selecting my work among all the works of the contestants. The experimental data is provided by subsidiary company of China mobile in Xinjiang. The research project comes from China Mobile Group Xinjiang Limited Company Development Found Projects (Projects No. XJM2012-01). Finally, thanks for the guidance of my teacher.

### REFERENCES

Cristianini, N. and J. Shawe-Taylor, 2004. Introduction to Support Vector Machines. Publishing House of Electronics Industry, Beijing.

- Hai-Shen, W. and C. Xiao-Ling, 2006. Power load forecasting with least squares support vector machines and chaos theory. Proceeding of the 6th World Congress on Intelligent Control and Automation. Dalian, pp: 4369-4373.
- Hua-Li, F. and L. Yuan, 2011. Network traffic prediction based on wavelet analysis and AR-LSSVM. *Comput. Eng. Appl.*, 47(20): 88-90.
- Hua, Y. and W. Hong, 2011. Study on time series forecasting based on least squares support vector machine. *Comput. Simul.*, 2(28): 225-227.
- Jie, L., L. Xian-Xing and H. Zhi-Jie, 2007. Research on the ARMA-based traffic prediction algorithm for wireless sensor network. *Electron. Inform. Technol.*, 5(29): 1224-1227.
- Jun, H., H. Yu-Qing and X. Zhong-Qing, 2008. Network traffic prediction models based on wavelet transform. *Comput. Eng.*, 34(19): 112-114.
- Jun-Song, W., 2009. Modeling and prediction for network traffic based on Elman neural network. *Comput. Eng.*, 35(9): 190-191.
- Rui, H. and J. Zhen-Hong, 2011. Telephone traffic load prediction based on SVR with DE-strategy. *Comput. Eng.*, 1(2): 178-180.
- Wang, H.J. and L. Shen, 2010. Adjustments based on wavelet transform ARIMA model for network traffic prediction. Proceeding of the International Conference on Computer Engineering and Technology. Chengdu, pp: 520-523.
- Xiao-Dong, W., Z. Hao-Ran and Z. Chang-Jiang, 2005. Prediction of chaotic time series using LS-SVM with automatic parameter selection. Proceeding of the 6th International Conference on Parallel and Distributed Computing, Application and Technologies, pp: 962-965.
- Xiao-Qiang, S. and M. Xian-Min, 2011. Modeling and simulation of chaotic time series prediction. *Comput. Simul.*, 5(28): 228-231.
- Xiao-Tian, C. and L. Jing-Xian, 2011. Network traffic prediction based on wavelet transformation and FARIIMA. *J. Commun.*, 4(32): 153-157.
- Xu-Qi, W. and Z. Zong-Tao, 2011. Prediction model of network traffic based on modified Elman neural network. *Microcomput. Appl.*, 12(12): 24-28.
- Zhi-Hui, Z., S. Yun-Lian and J. Yu, 2007. Short team load forecasting based on EMD and SVM. *High Voltage Eng.*, 33(5): 118-122.